WHAT IS CLAIMED IS:

1. A network data classifier configured to statistically classify data and comprising:

a network interface configured to receive packets carrying the data;

a feature extraction hardware block coupled to the network interface and configured to extract at least one feature from the received data;

a statistical classifier coupled to the feature extraction and configured to statistically classify the data in accordance with the at least one extracted feature; and

a policy engine coupled to the statistical classifier and configured to define a rule corresponding to the data class, wherein the statistical classifier is further configured to statistically classify the data at a same rate at which the network interface receives the packets.

2 . The network classifier of claim 1 wherein the rate at which the packets are received is greater than or equal to 100 Mbits/sec.

3. The network classifier of claim 1 further comprising:

a flow identifier coupled to the network interfaces and configured to identifying a flow to which each of the received packets belongs;

a flow assembler coupled to the flow identifier and configured to reorder the received packets such that the order of the reordered packets matches the order in which they were transmitted; and

a flow database configured to the flow assembler and configured to maintain a record for each identified flow.

4. The network classifier of claim 3 wherein the record for each identified flow includes at least one of an identification number, source and destination addresses of the received packets, protocol identification number, information used by the feature extraction hardware block and information used by the statistical classifier.

5. The network classifier of claim 4 further comprising:

a host interface configured to receive the packets from a host system.

6. The network classifier of claim 4 further comprising:

a host interface configured to receive the data from a host system.

7. The network classifier of claim 5 wherein the host interface is coupled to a device selected from a group consisting of microprocessor and network processor.

8. The network classifier of claim 7 wherein the host system is selected from a group consisting of firewall, router, switch, network appliance, security system, anti-virus system, anti-spam system, intrusion detection system, content filtering system, mail server, web server, quality of service provisioner, and gateway.

9. The network classifier of claim 8 wherein the host system is coupled to at least one of the flow identifier, the flow assembler, the feature extraction hardware block, the statistical classifier, and the flow database via one or more application programming interface.

10. The network classifier of claim 1 wherein the feature extractor is programmable.

11. The network classifier of claim 1 wherein the statistical classifier is programmable.

12. The network classifier of claim 1 wherein the policy engine is programmable.

13. The network classifier of claim 1 wherein the received data is one of messages, files, streams, documents, web pages, and e-mails.

14. The network classifier of claim 1 wherein the network interface is configured to interface with at least one of an Ethernet network, a SONET network, and an ATM network.

15. The network classifier of claim 1 wherein the packets are received via an Internet Protocol (IP) network.

16 . The network classifier of claim 1 wherein the feature extraction hardware block is configured to match extract features against a database of textual patterns.

17. The network classifier of claim 3 wherein the statistical classifier is configured to correlate events between one or more data flows

18. The network classifier of claim 11 wherein the statistical classifier includes at least one of linear discriminant classifier, artificial neural network classifier, support vector machine classifier, Bayesian network classifier, decision tree classifier; and nearest neighbor classifier.

19 . The network classifier of claim 18 wherein the artificial neural network classifier is configured to operate in accordance with an activation function selected from the group consisting of sigmoid function, hyperbolic tan function, Gaussian radial basis function, exponential radial basis function, and a non-linear function.

20. The network classifier of claim 18 wherein the support vector machine classifier is configured to operate in accordance with a kernel function selected from a group consisting of a linear projection function, polynomial function, piece-wise linear function, sigmoid function, Gaussian radial basis function, exponential radial basis function, and a non-linear transformation function.

21. The network classifier of claim 18 wherein the nearest neighbor classifier is configured to operate in accordance with a distance metric selected from a group consisting of Euclidean distance, Mahalanobis distance, and Manhattan distance.

22. The network classifier of claim 18 wherein the statistical classifier further generates a probability associated with a multitude of classes for the received data.

23. The network classifier of claim 22 wherein the statistical classifier classifies the received data for at least one of the applications selected from a group consisting of intrusion detection, content filtering, anti-spam, anti-virus, bandwidth management, quality of service provisioning, and network monitoring.

24. The network classifier of claim 1 wherein the at least one feature is selected from a group consisting of indicator vector, histogram, multitude of statistics associated with the data, mathematical transformation, timing information, and network events.

25. The network classifier of claim 3 wherein the feature extraction hardware block stores a history of the data it receives in the flow database, said history being used to extract the features from the received data.

26. The apparatus of claim 3 furthermore comprising:

a data flow multiplexer, the data flow multiplexer being coupled to the one or more of a plurality of network interfaces, the data flow multiplexer coupled to the one or more of a plurality of feature extraction devices, the data flow multiplexer providing for context switching between one or more of a plurality of data flows; and

a data flow context database, the data flow context database coupled to the data flow multiplexer, the data flow context database providing for retaining of state of said one or more of a plurality of data flows for said context switching.

27. The apparatus of claim 1, wherein said statistical classifier further comprises:

a lookup table configured to store weights for a multitude of events associated with the network data;

an adder coupled to add the weights it receives from the look-up table;

a register configured to store a value;

an accumulator; and

a multiplexer configured to deliver to the accumulator one of the added weights it receives from the adder at its first input terminal and the value it receives from the register at its second input terminal, the accumulator further configured to supply a summation of the added weights to the adder.

28. The integrated circuit of claim 27 furthermore comprising:

a hardware logic block configured to apply one of linear and non-linear functions to the summation stored in the accumulator.

29. The integrated circuit of claim 28 wherein the hardware logic block is configured to apply a non-linear function to the summation stored in the accumulator using lookup table.

30. The integrated circuit of claim 28 wherein the hardware logic block is formed in a programmable device.

31. The integrated circuit of claim 28 wherein the register is programmable.

32. The integrated circuit of claim 28 wherein the hardware logic block is programmable.

33. An integrated circuit configured to perform wire-speed computations for use in statistical classification of network data, the integrated circuit comprising:

a lookup table configured to store weights for a multitude of events associated with the network data;

an adder coupled to add the weights it receives from the look-up table;

a register configured to store a value;

an accumulator; and

a multiplexer configured to deliver to the accumulator one of the added weights it receives from the adder at its first input terminal and the value it receives from the register at its second input terminal, the accumulator further configured to supply a summation of the added weights to the adder.

34. The integrated circuit of claim 33 wherein said integrated circuit is a field programmable gate array.

35. The integrated circuit of claim 33 furthermore comprising:

a hardware logic block configured to apply a non-linear function to the summation stored in the accumulator.

36 . The integrated circuit of claim 35 wherein the hardware logic block is configured to apply a non-linear function to the summation stored in the accumulator using lookup table.

37. The integrated circuit of claim 35 wherein the hardware logic block is formed in a programmable device.

38. The integrated circuit of claim 35 wherein the register is programmable.

39. The integrated circuit of claim 35 wherein the hardware logic block is programmable.

40. A method for statistically classifying data, the method comprising:

receiving packets carrying the data;

extracting at least one feature from the received data;

statistically classifying the data in accordance with the at least one extracted feature and at a same rate at which the packets are received; and

applying a rule corresponding to the data class.

41. The method of claim 40 wherein the rate at which the packets are received is greater than or equal to 100 Mbits/sec.

42. The method of claim 40 further comprising:

identifying a flow to which each of the received packets belongs;

reordering the received packets such that the order of the reordered packets matches the order in which they were transmitted; and

maintaining a record for each identified flow.

43. The method of claim 42 wherein the record for each identified flow includes at least one of an identification number, source and destination addresses of the received packets, protocol identification number, information used for extracting the at least one feature extractor and information used to statistically classify the data.

44. The method of claim 43 further comprising:

receiving the packets from a host system.

45. The method of claim 43 further comprising:

receiving the data from a host system.

46. The method of claim 44 wherein the host system is selected from a group consisting of microprocessor and a network processor.

47. The method of claim 46 wherein the host system is selected from a group consisting of firewall, router, switch, network appliance, security system, anti-virus

system, anti-spam system, intrusion detection system, content filtering system, mail server, web server, quality of service provisioner, and gateway.

48. The method of claim 46 further comprising:
coupling the host system to one or more application programming interfaces.

49. The method of claim 40 wherein the received data is one of messages, files, streams, documents, web pages, and e-mails.

50. The method of claim 40 wherein the packets are received via one of an Ethernet network, a SONET network, and an ATM network.

51. The method of claim 40 wherein the packets are received via an Internet Protocol (IP) network.

52. The method of claim 40 further comprising:
matching the extract features against a database of textual patterns.

53. The method of claim 42 further comprising:
correlating events between one or more data flows.

54. The method of claim 53 wherein the statistically classifying of the data is carried out using a statistical classifier that includes at least one of linear discriminant classifier, artificial neural network classifier, support vector machine classifier, Bayesian network classifier, decision tree classifier; and nearest neighbor classifier.

55. The method of claim 54 wherein the artificial neural network classifier is configured to operate in accordance with an activation function selected from the group consisting of sigmoid function, hyperbolic tan function, Gaussian radial basis function, exponential radial basis function, and a non-linear function.

56. The method of claim 54 wherein the support vector machine classifier is configured to operate in accordance with a kernel function selected from a group consisting of a linear projection function, polynomial function, piece-wise linear function, sigmoid function, Gaussian radial basis function, exponential radial basis function, and a non-linear transformation function.

57. The method of claim 54 wherein the nearest neighbor classifier is configured to operate in accordance with a distance metric selected from a group consisting of Euclidean distance, Mahalanobis distance, and Manhattan distance.

58. The method of claim 54 wherein the statistical classifier further generates a probability associated with a multitude of classes for the received data.

59. The method of claim 58 wherein the statistical classifier classifies the received data for at least one of the applications selected from a group consisting of intrusion detection, content filtering, antivirus, bandwidth management, quality of service provisioning, anti-spam, and network management.

60. The method of claim 40 wherein the at least one feature is selected from a group consisting of indicator vector, histogram, multitude of statistics associated with the data, mathematical transformation, timing information, and network events.

61. The method of claim 42 further comprising:

stores a history of the received data, said history being used to extract the features from the received data.

62. The method of claim 42 further comprising:

multiplexing the data so as to provide for context switching between one or more of a plurality of data flows; and

retaining states of said one or more of a plurality of data flows for said context switching.